Article:
Sormaz, Mladen, Murphy, Charlotte, Wang, Hao Ting et al. (6 more authors) (2018) Default mode network can support the level of detail in experience during active task states. Proceedings of the National Academy of Sciences of the United States of America. pp. 9318-9323. ISSN 1091-6490

https://doi.org/10.1073/pnas.1721259115

Default mode network can support the level of detail in experience during active task states

Mladen Sormaz[a,1], Charlotte Murphy[a], Hao-ting Wang[a], Mark Hymers[a], Theodoros Karapanagiotidis[a], Giulia Poerio[b], Daniel S. Margulies[c], Elizabeth Jefferies[a], and Jonathan Smallwood[a,1]

[a]Department of Psychology, University of York, Heslington, York YO10 5DD, United Kingdom; [b]Department of Psychology, The University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom; and [c]Centre National de la Recherche Scientifique (CNRS) UMR 7225, Frontlab, Institut du Cerveau et de la Moelle Épinière, Paris, France

Edited by Marcus E. Raichle, Washington University in St. Louis, St. Louis, MO, and approved July 24, 2018 (received for review December 6, 2017)

Regions of transmodal cortex, in particular the default mode network (DMN), have historically been argued to serve functions unrelated to task performance, in part because of associations with naturally occurring periods of off-task thought. In contrast, contemporary views of the DMN suggest it plays an integrative role in cognition that emerges from its location at the top of a cortical hierarchy and its relative isolation from systems directly involved in perception and action. The combination of these topographical features may allow the DMN to support abstract representations derived from lower levels in the hierarchy and so reflect the broader cognitive landscape. To investigate these contrasting views of DMN function, we sampled experience as participants performed tasks varying in their working-memory load while inside an fMRI scanner. We used self-report data to establish dimensions of thought that describe levels of detail, the relationship to a task, the modality of thought, and its emotional qualities. We used representational similarity analysis to examine correspondences between patterns of neural activity and each dimension of thought. Our results were inconsistent with a task-negative view of DMN function. Distinctions between on- and off-task thought were associated with patterns of consistent neural activity in regions adjacent to unimodal cortex, including motor and premotor cortex. Detail in ongoing thought was associated with patterns of activity within the DMN during periods of working-memory maintenance. These results demonstrate a contribution of the DMN to ongoing cognition extending beyond task-unrelated processing that can include detailed experiences occurring under active task conditions.

default model network | principal gradient | mind wandering | representational similarity analysis

The functions of regions of transmodal cortex, such as those in the default mode network (DMN), remain a puzzle (1). They can deactivate during tasks (2, 3) and increase activity during periods of off-task thought (4–7)—leading initial accounts to emphasize task-negative functions for these regions (8, 9). An alternative perspective emerges from work showing regions in the DMN increase activity in tasks when cognition and behavior benefit from memory, such as when decisions depend on information from a prior trial (10, 11), when task-relevant stimuli are supported by long-term memory (12), when participants retrieve a task context from memory (13), or when participants have encoded the rule upon which their actions can be based (14). In addition, in conditions when activity in the DMN decreases, such as during more demanding semantic memory or working-memory tasks, connectivity between this network and task positive systems can increase, indicating that regions that make up this network continue to make an important contribution to cognition (15, 16). These recent studies show the DMN can play a role in task processing and provide support for alternative accounts that have explored the functions of this network from a topographical perspective. These studies show the cortex is organized along a macroscale gradient, anchored at one end by regions serving primary sensorimotor functions, and at the other by transmodal regions that make up the DMN (17, 18). This view suggests neural systems can be understood along a spectrum, from unimodal systems more directly involved in acting in the here and now to transmodal association cortex, which support neural operations less tethered to input (18). The DMN falls at the most extreme transmodal end of this spectrum (17), which may enable neural signals within it to integrate signals from other brain regions, providing top-down predictions to lower levels in the hierarchy (19, 20). This combination of recent empirical work, and advances in our understanding of the DMN in the larger cortical context, cast the functional role of this system in a different light. These integrative accounts of DMN function suggest its topographic location may enable it to encode relatively abstract representations by integrating signals from regions lower in the hierarchy. In this way, it is assumed to support neural contexts that reflect more integrated patterns of cognition, including those taking place during a task (17, 21, 22).

In light of this converging evidence for a role of the DMN in tasks states, we reevaluated the role of this system in ongoing thought. Our study used experience sampling to explore the traditional perspective that the DMN is important in the off-task state, as well as the alternative, that, given the DMN's role in memory-guided cognition, this system contributes to ongoing thought during the active maintenance of task-relevant information in working memory. Prior studies exploring the neural basis of patterns of ongoing cognition have combined experience sampling with measures of neural activity, such as functional

Significance

Accounts of the default mode network (DMN) as task negative are partly based on evidence for a role of this system in off-task thought. We revisited the evidence for this assumption in a study combining experience sampling with functional neuroimaging. Whether thoughts were related or unrelated to an ongoing task was associated with patterns of neural activity in regions adjacent to unimodal sensorimotor cortex. In contrast, during periods of working-memory maintenance, activity patterns in the DMN were associated with whether thoughts were detailed. These results demonstrate that activity within the DMN encodes information associated with ongoing cognition that goes beyond whether attention is directed to the task, including detailed experiences during active task states.

Author contributions: M.S. and H.-t.W. designed research; M.S., C.M., and G.P. performed research; M.S., C.M., M.H., and T.K. contributed new reagents/analytic tools; M.S. and J.S. analyzed data; M.S., D.S.M., E.J., and J.S. wrote the paper; and D.S.M. and E.J. contributed to development of themes in the writing of the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: mladen.sormaz@york.ac.uk or smallwoodjm@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1721259115/-/DCSupplemental.

Published online August 27, 2018.

magnetic resonance imaging (fMRI). These often find DMN activity increases when participants classify their thoughts as unrelated to the task being performed (4–7). However, recent investigations have revealed ongoing thought has a more complex structure than previously recognized (23–25). This structure can be captured by dimensional approaches, such as multidimensional experience sampling (MDES), which describes patterns in experience derived from multiple features of self-reported data (26). This technique reveals dimensions that are broadly task negative, such as off-task thoughts about the self in the future, as well as components describing emotions, or abstract properties of experience, such as its level of detail (26–28). In addition, prior studies combining fMRI with experience sampling often failed to manipulate the demands of the ongoing task, which may make it more difficult to identify any role the DMN plays in ongoing thoughts during active task states.

The current study combines self-reports in conjunction with fMRI, to reexamine the role of the DMN in ongoing cognition. We used representational similarity analysis (RSA) (29) to examine relationships between patterns of experience and neural activity. The choice of RSA was motivated by the sensitivity of this type of technique to covert psychological states (30, 31). RSA also has the advantage that it examines similarities between neural patterns and representational models, rather than absolute levels of activity, removing concerns about baseline activity levels within the DMN (32). In our case, it allowed us to understand the extent to which the self-reported dimensions describing ongoing thought are reflected in patterns of neural activity. To probe the potential role of the DMN in active task states, we used a task that manipulated whether information must be maintained in working memory. Task-negative accounts predict that neural patterns within the DMN should distinguish whether attention has shifted from the task in hand. Integrative accounts of the DMN, on the other hand, predict a role in cognition that reflects relatively abstract representations, and, importantly does not preclude a role in active task states.

Results

Our first analysis mapped the dimensional structure of the self-reported data recorded using MDES. We collected data in a large sample of participants (n = 199) who performed the task in a behavioral laboratory (lab) setting and a subset (n = 30) who also performed the same task within the scanner. Experience-sampling measures were taken twice in the fMRI scanner and three times in the behavioral laboratory. We used a set of questions developed in our prior work that have been validated by comparison with objective indices, including neural function, task performance, and pupilometry (23, 26–28). For both the laboratory and scanner, we collated reports on a trial-by-trial basis (*Methods*) and used principal components analysis (PCA) to decompose these data in two separate analyses. For both decompositions, scree plots indicated four components (*SI Appendix*, Fig. S1). The dimensions produced from the analysis of each dataset are presented as word clouds in Fig. 1*A*. The independent decomposition of different samples revealed broadly similar patterns of loading. One component (detail) describes thoughts that are detailed (highest loading "detailed"). A second component (task unrelated) reflects whether attention was directed to the task or toward other matters (low loadings on "task" and high loadings on "self," "person," and "future"). Component 3 (modality) reflects a dissociation between thoughts that were "images" and those that were "words." A fourth component (emotion) reflected the emotional quality of ongoing experience, describing variance from positive to negative. We compared the loadings from the solutions from the data recorded inside and outside the scanner, and found the highest similarity for the task-unrelated component, and the lowest for the emotion component. The correlation between patterns of experience recorded across the scanner in each task using the 30 individuals for whom we had both measures revealed significant correlations between both detail and modality

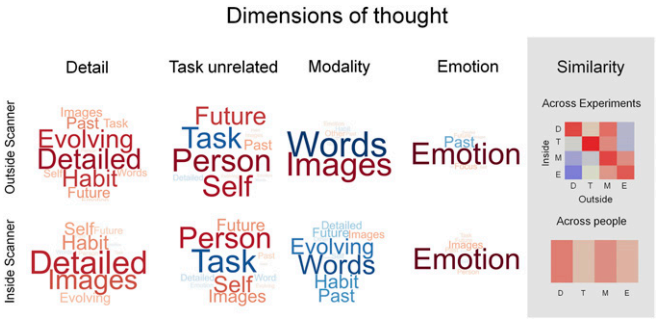

Fig. 1. Establishing components of ongoing thought using MDES. Word clouds describing the loadings on the four components of experience [detail (D), task unrelated (T), modality (M), and emotion (E)] derived from the data collected in this experiment. The font size describes the magnitude of the loading and the color describes the direction. Warm colors reflect positive loadings and cooler colors reflect negative loadings. On the *Right*, the *Upper* heat map describes the similarity between the structure of the component profiles across each session. The *Lower* heat map shows patterns of individual variation in the components in the laboratory and in the scanner; correlation values for these heat maps can be found in *SI Appendix*, Table S1.

(*SI Appendix*, Table S5). These analyses demonstrate the patterns of thought identified using MDES were stable, whether recorded inside or outside of the scanner, and were reasonably consistent when measured on multiple times in the same individual.

We recorded experience while participants performed blocks of an easier 0-back, and a more demanding 1-back task, similar to those used in our prior studies (Fig. 2*A*). In both conditions, participants observed a sequence of trials containing pairs of shapes (squares, circles, and triangles). During the response period, they made visuospatial decisions regarding the location of shapes. In the 0-back blocks, decisions were made using spatial information presented on screen at the moment when a response was made. In contrast, in the 1-back blocks decisions were made using spatial information from the prior trial. During the 1-back task, decisions depended on which side of the screen the shape was on during the previous trial. In the 1-back blocks, therefore, participants were required to maintain task-relevant information in working memory. A consequence of this manipulation is that when ongoing cognition is sampled in the 1-back task, it captures periods when task-relevant information is maintained in working memory, while the same period in the 0-back task does not require working-memory maintenance and is more conducive to thinking about task-unrelated matters (10, 27, 28). At intermittent intervals, when a target trial could have occurred, participants were instead asked to provide answers to the experience-sampling questions (*Methods*, Fig. 2*A*, and *SI Appendix*, Table S2).

We calculated inverse efficiency to characterize task performance (*Methods*) with a smaller score indicating better performance. Consistent with expectations, participants performed worse on the 1-back task: lab, $t(198) = 3.1$, $P = 0.002$, 1-back $M = 11.5$ (SE = 0.32), 0-back $M = 10.10$ (SE = 0.17); scanner, $t(29) = 2.43$, $P = 0.025$, 1-back $M = 9.08$ (SE = 0.1), 0-back $M = 8.69$ (SE = 0.2). Consistent with our prior work (10, 27, 28) ongoing cognition was more focused on information about social and self-relevant matters, and less about the task: lab, $t(198) = 10.6$, $P < 0.001$, 0-back $M = 0.03$ (SE = 0.04), 1-back $M = -0.07$ (SE = 0.05); scanner, $t(29) = 5.89$, $P < 0.001$, 0-back $M = 0.27$ (SE = 0.08), 1-back $M = -0.20$ (SE = 0.10). Inside the scanner, thoughts in the 1-back task were more often rated as in the form of images, $t(198) = 4.66$, $P < 0.001$, 0-back $M = 0.06$ (SE = 0.05), 1-back $M = -0.06$ (SE = 0.05), and more detailed, $t(29) = 2.24$, $P < 0.05$, 1-back $M = 0.10$ (SE = 0.13), 0-back $M = -0.01$ (SE = 0.13).

We also conducted a univariate analysis comparing the neural activity between the two tasks, focusing on the analysis window

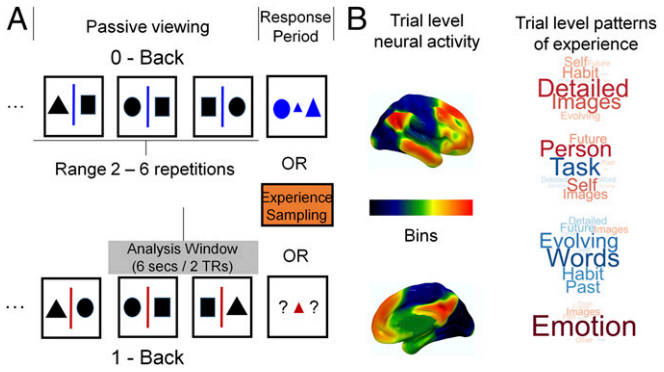

Fig. 2. Workflow for a representational similarity analysis examining links between patterns of neural response along the principle gradient and momentary changes in experience. (*A*) The task paradigm used to measure ongoing thought in the laboratory and the associated patterns of neural activity. Ongoing experience was measured in a task which alternated between an easy 0-back task and a more demanding 1-back task. For the purpose of analysis, we examined neural signals in the 6 s before the experience-sampling probe or the target. (*B*) Examining representational similarity at each point on the principle gradient. We used regions of interest that correspond to percentile steps along principal gradient from the decomposition performed by Margulies et al. (17). We used each mask as a region of interest analysis for a RSA comparing the common variance between neural patterns with the patterns of self-reported experience on a trial-by-trial basis. The word clouds represent the loadings of the different items for each of the patterns of experience.

indicated in Fig. 2. This allows a comparison of situations when task-relevant information must be maintained in working memory (1-back) and situations when it need not (0-back). Higher activity was observed in the 0-back task in regions of medial prefrontal cortex, posterior cingulate cortex/retrosplenial cortex, as well as regions of lateral/medial occipital cortex. Many of these correspond to regions within the DMN. In contrast, regions of left dorsolateral prefrontal cortex, inferior parietal sulcus, and posterior middle temporal gyrus, as well as the presupplementary motor area (bilaterally) were more active during the 1-back condition (*SI Appendix*, Fig. S3 and Table S1). Many of these are regions of the frontoparietal network (FPN). Greater activity in FPN regions during 1-back and higher DMN activity in 0-back were observed in a prior study using a similar paradigm (10).

Our primary goal was to assess the similarity between patterns of neural activity and patterns of ongoing thought recorded over the 6 s before the experience-sampling probe. Fig. 2*B* presents the workflow for this analysis. Following ref. 17 the brain was divided into 20 bins based on their percentile location along the principal gradient. These regions of interest were identified in the analysis performed in ref. 17. We conducted an RSA (29) in each region comparing variation in each of the dimensions identified by the PCA with neural activity (*Methods*). We assume neural activity over the analysis period contains information relevant to the experiential state as reported at the subsequent experience-sampling probe. In this context, RSA can determine where neural response patterns corresponds to the dimensions of ongoing thought determined using PCA. *SI Appendix*, Fig. S2 presents the distribution of values used in these analyses in the form of violin plots. Our regions of interest reflect percentiles along the principle gradient identified by ref. 17, so if regions of the DMN are important for a particular type of experience, then similarity between self-reports and neural responses should be higher close to the transmodal end of the gradient. After calculating the representational similarity for each region of the gradient for each dimension, we analyzed whether there were associations on the principal gradient between neural patterns and different dimensions of thought. We compared the model fit (correlation coefficient) of each component in each region to zero, using a sequence of one-tailed *t* tests. This determines if the observed association was significantly greater than zero. Before our analysis, we down-sampled the data to bins graded at every 10% to minimize the number of comparisons in our analysis. To control our false positive rate, we performed a Bonferroni correction, controlling for each of four dimensions, two tasks, and 10 bins. This yielded an alpha value of $P < 0.00125$ (one tailed).

Our analysis sought to determine whether regions along the principal gradient show associations between the patterns of neural response and aspects of experience (Figs. 3–5). During the 1-back task, the degree of task focus, and the level in detail for experiences were associated with neural patterns at specific positions along the principal gradient. The extent to which attention was disengaged from the task was associated with patterns in regions close to the unimodal end of the gradient, corresponding to regions of motor, premotor, and parietal cortex. This indicates regions outside the DMN provide information on whether patterns of thought are focused on a task or on other mental content (e.g., social or autobiographical information). In contrast, the detail in experience was predictable at an intermediate location along the gradient (percentiles 50–60, indicated in purple in Fig. 5), which included cortex within temporal parietal junction, posterior insula, and motor and visual cortex. Importantly, detailed thought was associated with activity in the most transmodal regions, corresponding to core DMN (posterior cingulate, lateral temporal cortex, angular gyrus, and medial prefrontal cortex, indicated in the palest colors in Fig. 5). These results indicate that neural patterns

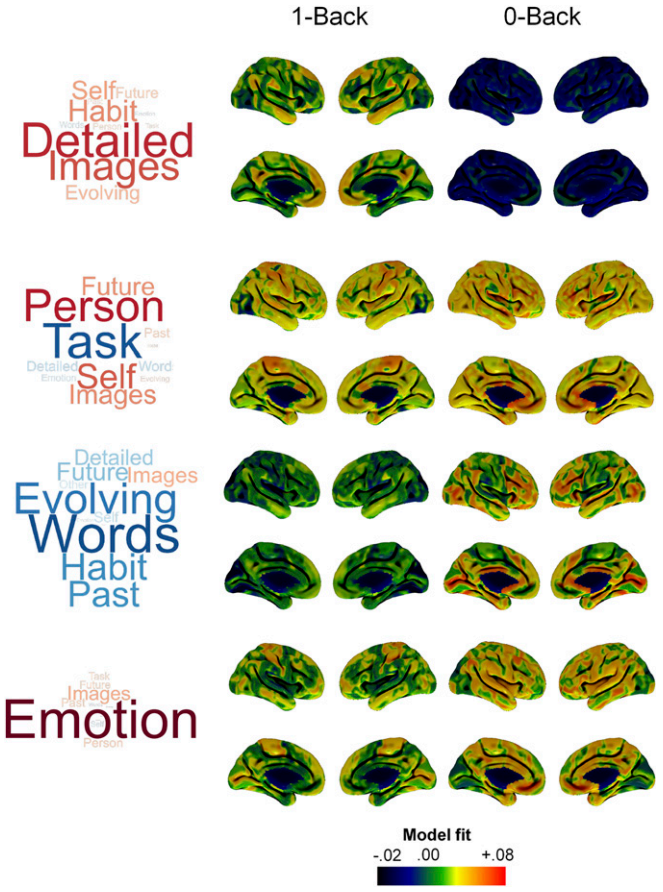

Fig. 3. The results of a RSA examining the association between neural activity and four dimensions of experience. The word clouds describe the loadings of items for each principal component, each describing a dimension of experience. The inflated brains show average RSA model correlation values, averaged across participants, for each bin of the principal gradient.

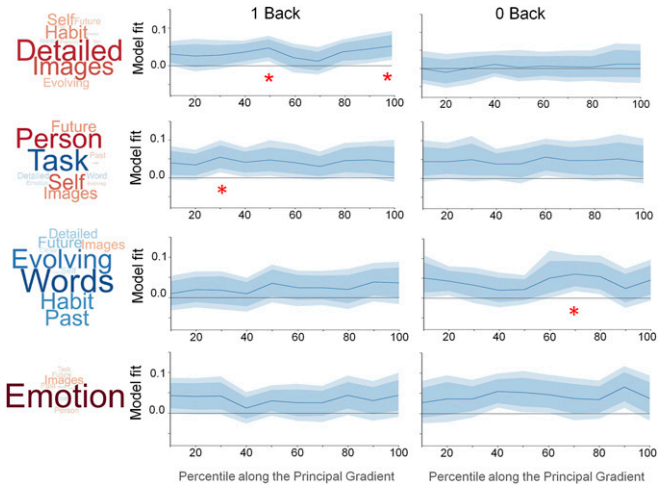

Fig. 4. Ribbon plots showing model fits plotted at each percentile of the principal gradient. The inner range of the confidence interval presents the 95% CI; the outer confidence interval describes the 99.5% CI. The regions indicated by the red asterisk describe regions that show a brain-experience association after applying the Bonferroni correction.

within regions of DMN are associated with the level of detail during working maintenance, suggesting a role in cognition extending beyond the off-task state. In the easier 0-back task, we found no associations with whether experience was directed to the task, whether it was detailed, or with the emotional tone of the experience. However, there was an association between the modality of experience and patterns of neural activity in the easier 0-back task in bin 7, including ventromedial prefrontal cortex, lateral prefrontal cortex, superior and middle temporal sulci, and regions in lateral parietal cortex (indicated in orange in Fig. 5).

Discussion

Our study reexamined the assumption that neural activity within the DMN reflects the difference between the presence of an on-task or an off-task state. We compared a task-negative view of the DMN with an alternative that it supports representations of task-relevant information in working memory. Although we could predict whether attention was directed toward or away from the task based on patterns of neural activity, the regions in which this was possible were located relatively close to unimodal sensorimotor cortex—including motor, premotor, and dorsal parietal regions. These brain areas serve functions allied to perception and action. It is possible, therefore, that this association represents a role for motor preparation in the on-task state during the 1-back task, an interpretation supported by prior studies linking the off-task state to poor motor control (33). We found that during periods of active working-memory maintenance, activity within the DMN was associated with the level of detail in ongoing thought. Together, these observations are problematic for task-negative accounts: regions outside of the DMN are diagnostic of whether attention is directed away from the task, and neural activity within this network was associated with detailed experiences during an active task state. Based on our results, patterns of activity within the DMN are neither necessary nor sufficient to determine if attention is directed away from the task.

While our study provides little support for task-negative accounts of the contribution of the DMN to ongoing thought, several features of our data are consistent with views of this system as integrating information from regions lower in the cortical hierarchy to form a representation of the ongoing neurocognitive context (14, 17, 19, 22). Our analysis divided the brain into a spectrum of regions of interest spanning those in unimodal regions of cortex (i.e., visual and motor cortex) to regions with more transmodal functions, the most extreme of which are those in the DMN. In both regions adjacent to unimodal cortex, and those furthest along this spectrum, neural activity encoded patterns of detailed thought during periods of working-memory maintenance. These data are consistent with a view of DMN function that enables detailed cognitive representations through interactions with regions lower in a cortical hierarchy. Importantly, a role for regions in the DMN in detailed experiences during working memory is consistent with graph theoretical analyses, showing a role for this system in working-memory processes (16), and from studies that show that elements of this network are implicated in specificity (34) or detail (35) during memory retrieval. It is also consistent with links between individual variation in detailed thought to the functional connectivity of hippocampus with posterior cingulate cortex (26).

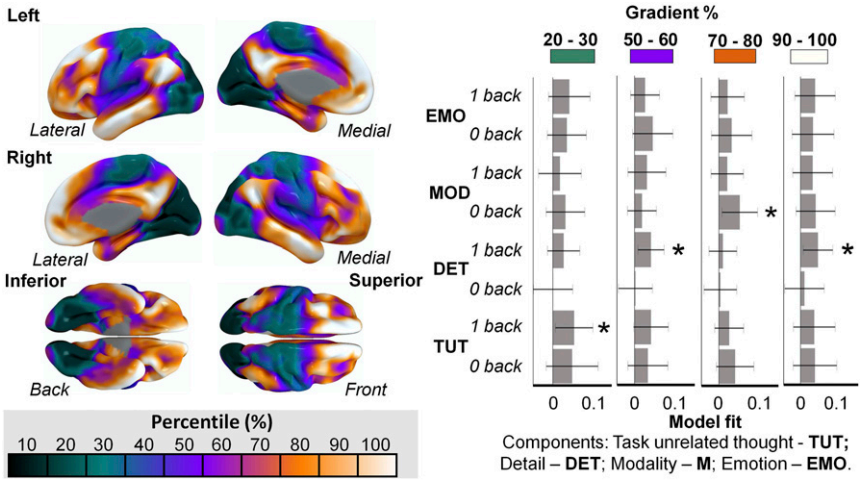

Fig. 5. Distribution of significant model fits for each experience displayed in the context of the principle gradient. The inflated brains on the *Left* display the continuous distribution of the gradient values across the cortical surface down-sampled into the percentile bins used in our analyses. The *Right* displays the specific percentiles on the gradient in which associations with each component for one task were identified. The bar graphs show the average correlation between each component in each task condition in each region of the gradient. The error bars show the 99.5% confidence intervals. The asterisks indicate situations where the model fit was significantly greater than zero, following Bonferroni correction.

Our study also found patterns of neural activity in regions intermediate between unimodal and transmodal regions that distinguish the modality in which ongoing thought unfolds in the 0-back task. Whether experience was based on verbal or visual codes was associated with neural patterns at around the 70th percentile of the principal gradient, marked in orange in Fig. 5. It is possible that we were able to classify the modality of ongoing thought in the easier 0-back task because under these conditions, experience is relatively unconstrained by input. In contrast, in the 1-back condition, participants must encode and maintain visual information, reducing trial by trial variation in the modality of ongoing thought, and limiting the ability to identify associations between experience and neural activity. Notably, within the temporal lobe the regions of high classification accuracy corresponded to the sulci that surround the middle temporal gyrus, and, in parietal cortex, to regions that form the boundary around the angular gyrus. Regions around the 70th percentile of the gradient encompass some of the regions identified by our prior task-based study as being sensitive to the modality in which a stimulus was presented (auditory vs. visual), but not its meaning (36). A common relationship between task-relevant differences in modality and dissociations between verbal and visual codes of ongoing thought provides confidence in our current data. It also provides converging evidence for contemporary accounts of semantic processing (37–39). These assume regions in the temporal lobe support the progressive integration of neural signal from peripheral regions on the lateral surface into a central core that represents information in a more amodal manner. In this regard, we note that in our current data, regions of cortex where neural patterns were associated with differences in the modality of ongoing experience were high, associations with the level of detail were low (Fig. 4). This partial dissociation supports the view that within the temporal lobe, modality-specific information is represented in more peripheral regions than are more abstract descriptions of experience (such as its level of detail), which are represented in the transmodal core.

In closing, it is worth considering why prior results linked activity in the DMN to off-task thought (4–7). One possibility is that this pattern emerged because these studies tend to use simple tasks that often do not vary task demands within an individual. The absence of such a task manipulation can make identification of patterns of experience linked to more active task states challenging (although see ref. 7). In addition, they often measure experience using only a small number of questions, and this neglects the heterogeneous nature of ongoing experience (40). Off-task thought differs from periods of external task focus in both the process it entails and the content it involves. For example, our data suggest off-task thoughts include episodic and social information, because they load on terms like "person" and "future." Prior studies have shown that off-task thoughts depend more upon memory and less upon sensory processing (41) and they can also have close ties to affect (42, 43). Simplistic binary classifications of experience into on/off-task states, or any other dichotomous system, will necessarily confound overlapping aspects of experience. Our dimensional approach, in contrast, allows multiple distinct aspects of experience to be identified and shows neural activity in the DMN is not synonymous with the off-task state.

Methods

Participants. A total of 207 participants (females = 132, age = 20.2, SD = 2.35 y). Eight participants were excluded from the final analysis for failing to complete all sessions. A total of 34 participants participated in the scanning session (16 females). Four participants were excluded from the final analysis (2 fell asleep, 1 withdrew from the second session, and 1 showed movement >1 mm in more than 50% of runs). The study was carried out in accordance within the guidelines of, and with ethical approval from the York Neuroimaging Centre and the University of York's Psychology Department. All participants gave informed consent.

MDES. Experiential content was measured using MDES with a battery of 13 questions (*SI Appendix*, Table S2). In the scanner, probes occurred six times per run. Participants had 5 s to respond to each question and in total the probe question points were of a fixed 65-s duration. In each run, there were an average of 3 thought probes in the 0-back condition and 3 in the 1-back condition. In total, we administered a total of 48 probes in the scanner for each person. In each session (4 fMRI runs × 6 thought probes), an average of 24 (SD = 3.30) MDES probes occurred; in the 0-back condition, an average of 12 (SD = 2.36); and in the 1-back condition, an average of 12 (SD = 2.24). In the laboratory, MDES probes occurred on a quasirandom basis to minimize the likelihood of anticipating the probes. There was a 20% chance of a MDES probe instead of a target with a minimum of 1 probe per condition.

MRI Acquisition. Structural and functional data were acquired using a 3T GE HDx Excite MRI scanner with an eight-channel phased array head coil (GE) tuned to 127.4 MHz, at the York Neuroimaging Centre. Structural MRI acquisition was based on a T1-weighted 3D fast spoiled gradient echo sequence [repetition time (TR) = 7.8 s, echo time (TE) = minimum full, flip angle = 20°, matrix size = 256 × 256, 176 slices, voxel size = 1.13 × 1.13 × 1 mm]. Functional data were recorded using single-shot 2D gradient echo planar imaging (TR = 3 s, TE = minimum full, flip angle = 90°, matrix size = 64 × 64, 60 slices, voxel size = 3 × 3 × 3 mm^3, 180 volumes). A FLAIR scan with the same orientation as the functional scans was collected to improve coregistration between scans.

Data Analysis.

Task performance of 0-back and 1-back. In both fMRI and lab tasks, we recorded mean accuracy and reaction time (RT) for participants in the 0-back and 1-back experimental conditions (*SI Appendix*, Table S6). We calculated an inverse efficiency score (RT in milliseconds/accuracy in percent correct) for each participant in which small scores indicate better efficiency/performance (44).

PCA of MDES data. In both experiments, we created a single n-by-n matrix in which each row was a MDES probe and each column was a question. We used PCA with varimax rotation in SPSS version 24 to reduce the dimensionality of these data. Scree plots showed that extracting four principal components was appropriate in both cases (*SI Appendix*, Fig. S3). The resulting data were used as regressors of interest at the trial level in the imaging analysis. We also correlated PCA loadings across task environment (laboratory or scanner) to test their consistency (Fig. 1). Components were named based on the highest loading if the component was anchored at one end (i.e., detail). If they were anchored by two opposing scores (words or images), we named them based on the distinction these loadings imply (i.e., modality).

fMRI Analysis. Functional and structural data were preprocessed and analyzed using FMRIB's Software Library (FSL version 4.1, fsl.fmrib.ox.ac.uk/fsl/fslwiki/FEAT/) (45). Individual FLAIR and T1-weighted structural brain images were extracted using brain extraction tool (BET). Structural images were registered to the MNI-152 template using FMRIB's linear image registration tool (FLIRT). The functional data were preprocessed and analyzed using the FMRI expert analysis tool (FEAT). Individual subject analysis involved: motion correction using MCFLIRT; slice-timing correction using Fourier space time-series phase shifting; spatial smoothing using a Gaussian kernel of full width at half minimum (FWHM) 6 mm; grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 100 s). Although not always standard in RSA analysis, we chose to smooth our data for this analysis to increase the signal-to-noise ratio for each neural response pattern. This was motivated by the trial-by-trial nature of our comparisons which necessarily meant that the data were noisy. Consistent with prior studies (e.g., ref. 46) we found lower RSA fit correlations with the nonsmoothed data. We report the results of the nonsmoothed analysis in *SI Appendix*, Fig. S4.

First level analyses modeled six explanatory variables (EVs). EV 1 and 2 modeled time periods in which participants completed the 0-back task or 1-back condition. EV 3–6 modeled the four extracted principal components by assigning the beta weight of each extracted principal component to each thought at the trial level. PCA loadings were modeled in a 6-s time window (the minimum time period between thought probes or targets, labeled as the analysis window in Fig. 2). Individual participant data were entered into a higher-level fixed-effect analysis to measure and average neural response to the six EV's across all eight functional runs. Following this, for the univariate task comparison analysis a fixed-effects analysis was performed (FLAME, https://fsl.fmrib.ox.ac.uk/fsl/fslwiki). We defined task-specific neural responses by contrasting z-stat maps for 0-B > 1-B and 1-B > 0-B conditions. These maps were then registered to a high-resolution

T1-anatomical image and then onto the standard Montreal Neurological Institute (MNI) brain (MNI 152 average template). For these contrasts, following ref. 47, we first applied a cluster-forming threshold of $Z > 3.1$ to our analyses. This corresponds to a P value of >0.01. The resulting spatial maps were whole brain corrected at $P < 0.05$ FWE.

RSA. We used a version of RSA (29) to compare patterns of dissimilarity between neural and experiential data in each scanning session. To create neural representational dissimilarity matrices (RDMs), the fMRI data were modeled in the same way as described in the previous section except individual thought probe trials were not weighted by principal component weights.

Regions of interest for RSA. We used the same masks as in ref. 17. Regions of interest masks were created by projecting the cortical surface principal gradient map to the 2-mm volumetric MNI152 standard space (available at https://neurovault.org/images/24346/). The volumetric map was then binned into five-percentile increments and binarized. These are available here: https://neurovault.org/collections/3192/.

Neural RDMs. Neural RDMs were created separately for the 0-back and 1-back tasks by first correlating the neural response patterns between each pairwise combination of thought probes (using pyMVPA version 2), we then converted these into distance measures ($1 - r$ value) as recommended by Kriegeskorte et al. (29). Each participant had two neural RDMs per session (one each for 0-back tasks and 1-back tasks). These RDMs were created based on neural patterns in each individual's native echo planar imaging (EPI) brain space for each of the 20-gradient bins defined by ref. 17.

Principal component RDMs. Principal component RDMs were created in a similar fashion to the neural RDMs except that they were composed of the representational distances between principal component loadings for each pairwise combination of thought probes. This was done separately for each principal component [detail, task unrelated thought (TUT), modality, and emotion] in the 0- and 1-back conditions. The use of varimax rotation in the MDES decomposition allowed us to construct four separate models in each task that were not collinear.

RSA model comparison. We removed the diagonal components of each RDM and used one-half of the correlation matrix (the duplicated values). We used Spearman rank correlations (29) between each principal component RDM model (detail, TUT, modality, and emotion) in each condition (0-back, 1-back) and each neural RDM in each condition (0- and 1-back thought probes). RDMs were calculated for each participant and averaged across session, leaving eight correlation values for each participant: one for each principal component RDM compared with the neural RDM in that condition. The r values were Fisher transformed to ensure they fit assumption of normality for significance testing. Mean RSA model fit scores for each of the 20 bins were z scored and superimposed onto the inflated surface of a standard MNI152 brain in Fig. 4.

ACKNOWLEDGMENTS. This study was supported by awards (to J.S.) from the European Research Council (Wandering Minds 646927) and the Volkswagen Foundation (Wandering Minds 89440 and 89439), and the John Templeton Foundation, "Prospective Psychology Stage 2: A Research Competition" and Grant FLEXSEM 771863 (to E.J.).

1. Raichle ME (2015) The brain's default mode network. *Annu Rev Neurosci* 38:433–447.
2. Gusnard DA, Raichle ME, Raichle ME (2001) Searching for a baseline: Functional imaging and the resting human brain. *Nat Rev Neurosci* 2:685–694.
3. Raichle ME, et al. (2001) A default mode of brain function. *Proc Natl Acad Sci USA* 98: 676–682.
4. Stawarczyk D, Majerus S, Maquet P, D'Argembeau A (2011) Neural correlates of ongoing conscious experience: Both task-unrelatedness and stimulus-independence are related to default network activity. *PLoS One* 6:e16997.
5. Christoff K, Gordon AM, Smallwood J, Smith R, Schooler JW (2009) Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc Natl Acad Sci USA* 106:8719–8724.
6. Allen M, et al. (2013) The balanced mind: The variability of task-unrelated thoughts predicts error monitoring. *Front Hum Neurosci* 7:743.
7. Kucyi A, Esterman M, Riley CS, Valera EM (2016) Spontaneous default network activity reflects behavioral variability independent of mind-wandering. *Proc Natl Acad Sci USA* 113:13899–13904.
8. Anticevic A, et al. (2012) The role of default network deactivation in cognition and disease. *Trends Cogn Sci* 16:584–592.
9. Fox MD, et al. (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* 102:9673–9678.
10. Konishi M, McLaren DG, Engen H, Smallwood J (2015) Shaped by the past: The default mode network supports cognition that is independent of immediate perceptual input. *PLoS One* 10:e0132209.
11. Murphy C, et al. (2018) Distant from input: Evidence of regions within the default mode network supporting perceptually-decoupled and conceptually-guided cognition. *Neuroimage* 171:393–401.
12. Spreng RN, et al. (2014) Goal-congruent default network activity facilitates cognitive control. *J Neurosci* 34:14108–14114.
13. Crittenden BM, Mitchell DJ, Duncan J (2015) Recruitment of the default mode network during a demanding act of executive control. *eLife* 4:e06481.
14. Vatansever D, Menon DK, Stamatakis EA (2017) Default mode contributions to automated information processing. *Proc Natl Acad Sci USA* 114:12821–12826.
15. Krieger-Redwood K, et al. (2016) Down but not out in posterior cingulate cortex: Deactivation yet functional coupling with prefrontal cortex during demanding semantic cognition. *Neuroimage* 141:366–377.
16. Vatansever D, Menon DK, Manktelow AE, Sahakian BJ, Stamatakis EA (2015) Default mode network connectivity during task execution. *Neuroimage* 122:96–104.
17. Margulies DS, et al. (2016) Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc Natl Acad Sci USA* 113:12574–12579.
18. Buckner RL, Krienen FM (2013) The evolution of distributed association networks in the human brain. *Trends Cogn Sci* 17:648–665.
19. Braga RM, Sharp DJ, Leeson C, Wise RJ, Leech R (2013) Echoes of the brain within default mode, association, and heteromodal cortices. *J Neurosci* 33:14031–14039.
20. Leech R, Sharp DJ (2014) The role of the posterior cingulate cortex in cognition and disease. *Brain* 137:12–32.
21. Mesulam MM (1998) From sensation to cognition. *Brain* 121:1013–1052.
22. Braga RM, Leech R (May 6, 2015) Echoes of the brain: Local-scale representation of whole-brain functional networks within transmodal cortex. *Neuroscientist,* 1073858415585730.
23. Ruby FJ, Smallwood J, Engen H, Singer T (2013) How self-generated thought shapes mood: The relation between mind-wandering and mood depends on the socio-temporal content of thoughts. *PLoS One* 8:e77554.
24. Seli P, Ralph BCW, Konishi M, Smilek D, Schacter DL (2017) What did you have in mind? Examining the content of intentional and unintentional types of mind wandering. *Conscious Cogn* 51:149–156.
25. Christoff K, Irving ZC, Fox KC, Spreng RN, Andrews-Hanna JR (2016) Mind-wandering as spontaneous thought: A dynamic framework. *Nat Rev Neurosci* 17:718–731.
26. Smallwood J, et al. (2016) Representing representation: Integration between the temporal lobe and the posterior cingulate influences the content and form of spontaneous thought. *PLoS One* 11:e0152272.
27. Konishi M, Brown K, Battaglini L, Smallwood J (2017) When attention wanders: Pupillometric signatures of fluctuations in external attention. *Cognition* 168:16–26.
28. Medea B, et al. (July 21, 2016) How do we decide what to do? Resting-state connectivity patterns and components of self-generated thought linked to the development of more concrete personal goals. *Exp Brain Res,* 10.1007/s00221-016-4729-y.
29. Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis–Connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4.
30. Woo C-W, et al. (2014) Separate neural representations for physical pain and social rejection. *Nat Commun* 5:5380.
31. Horikawa T, Tamaki M, Miyawaki Y, Kamitani Y (2013) Neural decoding of visual imagery during sleep. *Science* 340:639–642.
32. Passow S, et al. (2015) Default-mode network functional connectivity is closely related to metabolic activity. *Hum Brain Mapp* 36:2027–2038.
33. Smallwood J, Beach E, Schooler JW, Handy TC (2008) Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *J Cogn Neurosci* 20:458–469.
34. Davey J, et al. (2015) Automatic and controlled semantic retrieval: TMS reveals distinct contributions of posterior middle temporal gyrus and angular gyrus. *J Neurosci* 35: 15230–15239.
35. Richter FR, Cooper RA, Bays PM, Simons JS (2016) Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *eLife* 5:e18260.
36. Murphy C, et al. (2017) Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. *Neuroimage* 147:19–31.
37. Ralph MA, Jefferies E, Patterson K, Rogers TT (2017) The neural and computational bases of semantic cognition. *Nat Rev Neurosci* 18:42–55.
38. Visser M, Embleton KV, Jefferies E, Parker GJ, Ralph MA (2010) The inferior, anterior temporal lobes and semantic memory clarified: Novel evidence from distortion-corrected fMRI. *Neuropsychologia* 48:1689–1696.
39. Visser M, Jefferies E, Lambon Ralph MA (2010) Semantic processing in the anterior temporal lobes: A meta-analysis of the functional neuroimaging literature. *J Cogn Neurosci* 22:1083–1094.
40. Smallwood J, Andrews-Hanna J (2013) Not all minds that wander are lost: The importance of a balanced perspective on the mind-wandering state. *Front Psychol* 4:441.
41. Poerio GL, et al. (2017) The role of the default mode network in component processes underlying the wandering mind. *Soc Cogn Affect Neurosci* 12:1047–1062.
42. Killingsworth MA, Gilbert DT (2010) A wandering mind is an unhappy mind. *Science* 330:932.
43. Poerio GL, Totterdell P, Miles E (2013) Mind-wandering and negative mood: Does one thing really lead to another? *Conscious Cogn* 22:1412–1421.
44. Townsend JT, Ashby FG (1983) *Stochastic Modeling of Elementary Psychological Processes* (CUP Archive, Cambridge).
45. Smith SM, et al. (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(Suppl 1):S208–S219.
46. Hendriks MHA, Daniels N, Pegado F, Op de Beeck HP (2017) The effect of spatial smoothing on representational similarity in a simple motor paradigm. *Front Neurol* 8:222.
47. Eklund A, Nichols TE, Knutsson H (2016) Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci USA* 113: 7900–7905.